0 3201090 #7

ENTERED   OIPE

RAW SEQUENCE LISTING   DATE: 04/18/2002 P5
PATENT APPLICATION: US/10/044,303   TIME: 11:06:15

Input Set : A:\Vectorsl.app
Output Set: N:\CRF3\04182002\J044303.raw

```
   3 <110> APPLICANT: Max-Planck-Gesellschaft e.V.
   5 <120> TITLE OF INVENTION: Protein expression and structure solution using
   6       specific fusion vectors
   8 <130> FILE REFERENCE: ST010209-EPA
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/044,303
C--> 11 <141> CURRENT FILING DATE: 2002-01-11
  13 <160> NUMBER OF SEQ ID NOS: 3
  15 <170> SOFTWARE: PatentIn Ver. 2.1
  17 <210> SEQ ID NO: 1
  18 <211> LENGTH: 765
  19 <212> TYPE: PRT
  20 <213> ORGANISM: Artificial Sequence
  22 <220> FEATURE:
  23 <223> OTHER INFORMATION: Description of Artificial Sequence: Partial myosin
  24       sequence of Dictyostelium; Component (1) of the
  25       recombinant protein M761-2R R238E
  27 <400> SEQUENCE: 1
  28 Met Asp Gly Thr Glu Asp Pro Ile His Asp Arg Thr Ser Asp Tyr His
  29  1               5                  10                  15
  31 Lys Tyr Leu Lys Val Lys Gln Gly Asp Ser Asp Leu Phe Lys Leu Thr
  32              20                  25                  30
  34 Val Ser Asp Lys Arg Tyr Ile Trp Tyr Asn Pro Asp Pro Lys Glu Arg
  35          35                  40                  45
  37 Asp Ser Tyr Glu Cys Gly Glu Ile Val Ser Glu Thr Ser Asp Ser Phe
  38          50                  55                  60
  40 Thr Phe Lys Thr Val Asp Gly Gln Asp Arg Gln Val Lys Lys Asp Asp
  41 65                  70                  75                  80
  43 Ala Asn Gln Arg Asn Pro Ile Lys Phe Asp Gly Val Glu Asp Met Ser
  44              85                  90                  95
  46 Glu Leu Ser Tyr Leu Asn Glu Pro Ala Val Phe His Asn Leu Arg Val
  47              100                 105                 110
  49 Arg Tyr Asn Gln Asp Leu Ile Tyr Thr Tyr Ser Gly Leu Phe Leu Val
  50          115                 120                 125
  52 Ala Val Asn Pro Phe Lys Arg Ile Pro Ile Tyr Thr Gln Glu Met Val
  53          130                 135                 140
  55 Asp Ile Phe Lys Gly Arg Arg Arg Asn Glu Val Ala Pro His Ile Phe
  56 145                 150                 155                 160
  58 Ala Ile Ser Asp Val Ala Tyr Arg Ser Met Leu Asp Asp Arg Gln Asn
  59              165                 170                 175
  61 Gln Ser Leu Leu Ile Thr Gly Glu Ser Gly Ala Gly Lys Thr Glu Asn
  62              180                 185                 190
  64 Thr Lys Lys Val Ile Gln Tyr Leu Ala Ser Val Ala Gly Arg Asn Gln
  65              195                 200                 205
```

```
 67 Ala Asn Gly Ser Gly Val Leu Glu Gln Gln Ile Leu Gln Ala Asn Pro
 68     210             215                 220
 70 Ile Leu Glu Ala Phe Gly Asn Ala Lys Thr Thr Arg Asn Asn Asn Ser
 71 225             230                 235                 240
 73 Ser Arg Phe Gly Lys Phe Ile Glu Ile Gln Phe Asn Ser Ala Gly Phe
 74             245                 250                 255
 76 Ile Ser Gly Ala Ser Ile Gln Ser Tyr Leu Leu Glu Lys Ser Arg Val
 77             260                 265                 270
 79 Val Phe Gln Ser Glu Thr Glu Arg Asn Tyr His Ile Phe Tyr Gln Leu
 80         275                 280                 285
 82 Leu Ala Gly Ala Thr Ala Glu Glu Lys Lys Ala Leu His Leu Ala Gly
 83         290                 295                 300
 85 Pro Glu Ser Phe Asn Tyr Leu Asn Gln Ser Gly Cys Val Asp Ile Lys
 86 305             310                 315                 320
 88 Gly Val Ser Asp Ser Glu Glu Phe Lys Ile Thr Arg Gln Ala Met Asp
 89             325                 330                 335
 91 Ile Val Gly Phe Ser Gln Glu Glu Gln Met Ser Ile Phe Lys Ile Ile
 92             340                 345                 350
 94 Ala Gly Ile Leu His Leu Gly Asn Ile Lys Phe Glu Lys Gly Ala Gly
 95             355                 360                 365
 97 Glu Gly Ala Val Leu Lys Asp Lys Thr Ala Leu Asn Ala Ala Ser Thr
 98         370                 375                 380
100 Val Phe Gly Val Asn Pro Ser Val Leu Glu Lys Ala Leu Met Glu Pro
101 385                 390                 395                 400
103 Arg Ile Leu Ala Gly Arg Asp Leu Val Ala Gln His Leu Asn Val Glu
104             405                 410                 415
106 Lys Ser Ser Ser Ser Arg Asp Ala Leu Val Lys Ala Leu Tyr Gly Arg
107             420                 425                 430
109 Leu Phe Leu Trp Leu Val Lys Lys Ile Asn Asn Val Leu Cys Gln Glu
110         435                 440                 445
112 Arg Lys Ala Tyr Phe Ile Gly Val Leu Asp Ile Ser Gly Phe Glu Ile
113         450                 455                 460
115 Phe Lys Val Asn Ser Phe Glu Gln Leu Cys Ile Asn Tyr Thr Asn Glu
116 465                 470                 475                 480
118 Lys Leu Gln Gln Phe Phe Asn His His Met Phe Lys Leu Glu Gln Glu
119             485                 490                 495
121 Glu Tyr Leu Lys Glu Lys Ile Asn Trp Thr Phe Ile Asp Phe Gly Leu
122         500                 505                 510
124 Asp Ser Gln Ala Thr Ile Asp Leu Ile Asp Gly Arg Gln Pro Pro Gly
125         515                 520                 525
127 Ile Leu Ala Leu Leu Asp Glu Gln Ser Val Phe Pro Asn Ala Thr Asp
128         530                 535                 540
130 Asn Thr Leu Ile Thr Lys Leu His Ser His Phe Ser Lys Lys Asn Ala
131 545                 550                 555                 560
133 Lys Tyr Glu Glu Pro Arg Phe Ser Lys Thr Glu Phe Gly Val Thr His
134             565                 570                 575
136 Tyr Ala Gly Gln Val Met Tyr Glu Ile Gln Asp Trp Leu Glu Lys Asn
137         580                 585                 590
139 Lys Asp Pro Leu Gln Gln Asp Leu Glu Leu Cys Phe Lys Asp Ser Ser
```

Input Set : A:\Vectors1.app
Output Set: N:\CRF3\04182002\J044303.raw

```
140        595                600                605
142 Asp Asn Val Val Thr Lys Leu Phe Asn Asp Pro Asn Ile Ala Ser Arg
143     610                615                620
145 Ala Lys Lys Gly Ala Asn Phe Ile Thr Val Ala Ala Gln Tyr Lys Glu
146 625                630                635                640
148 Gln Leu Ala Ser Leu Met Ala Thr Leu Glu Thr Thr Asn Pro His Phe
149                645                650                655
151 Val Arg Cys Ile Ile Pro Asn Asn Lys Gln Leu Pro Ala Lys Leu Glu
152         660                665                670
154 Asp Lys Val Val Leu Asp Gln Leu Arg Cys Asn Gly Val Leu Glu Gly
155         675                680                685
157 Ile Arg Ile Thr Arg Lys Gly Phe Pro Asn Arg Ile Ile Tyr Ala Asp
158         690                695                700
160 Phe Val Lys Arg Tyr Tyr Leu Leu Ala Pro Asn Val Pro Arg Asp Ala
161 705                710                715                720
163 Glu Asp Ser Gln Lys Ala Thr Asp Ala Val Leu Lys His Leu Asn Ile
164                725                730                735
166 Asp Pro Glu Gln Tyr Arg Phe Gly Ile Thr Lys Ile Phe Phe Arg Ala
167                740                745                750
169 Gly Gln Leu Ala Arg Ile Glu Glu Ala Arg Glu Gln Arg
170         755                760                765
```

173 <210> SEQ ID NO: 2
174 <211> LENGTH: 1016
175 <212> TYPE: PRT
176 <213> ORGANISM: Artificial Sequence
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Description of Artificial Sequence: Whole sequence
180   of recombinant protein M761-2R R238 E
182 <220> FEATURE:
183 <223> OTHER INFORMATION: The protein comprises as component (1) the aa
184   sequence of myosin II motor domain of
185   Dictyostelium, a three aa linker region and the
186   a-actinin aa sequence
188 <400> SEQUENCE: 2

```
189 Met Asp Gly Thr Glu Asp Pro Ile His Asp Arg Thr Ser Asp Tyr His
190    1            5                10                15
192 Lys Tyr Leu Lys Val Lys Gln Gly Asp Ser Asp Leu Phe Lys Leu Thr
193                20                25                30
195 Val Ser Asp Lys Arg Tyr Ile Trp Tyr Asn Pro Asp Pro Lys Glu Arg
196         35                40                45
198 Asp Ser Tyr Glu Cys Gly Glu Ile Val Ser Glu Thr Ser Asp Ser Phe
199         50                55                60
201 Thr Phe Lys Thr Val Asp Gly Gln Asp Arg Gln Val Lys Lys Asp Asp
202 65                70                75                80
204 Ala Asn Gln Arg Asn Pro Ile Lys Phe Asp Gly Val Glu Asp Met Ser
205                85                90                95
207 Glu Leu Ser Tyr Leu Asn Glu Pro Ala Val Phe His Asn Leu Arg Val
208                100                105                110
210 Arg Tyr Asn Gln Asp Leu Ile Tyr Thr Tyr Ser Gly Leu Phe Leu Val
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/044,303

DATE: 04/18/2002

TIME: 11:06:15

Input Set : **A:\Vectors1.app**

Output Set: **N:\CRF3\04182002\J044303.raw**

```
211         115                  120                  125
213 Ala Val Asn Pro Phe Lys Arg Ile Pro Ile Tyr Thr Gln Glu Met Val
214     130              135              140
216 Asp Ile Phe Lys Gly Arg Arg Arg Asn Glu Val Ala Pro His Ile Phe
217 145              150              155              160
219 Ala Ile Ser Asp Val Ala Tyr Arg Ser Met Leu Asp Asp Arg Gln Asn
220             165              170              175
222 Gln Ser Leu Leu Ile Thr Gly Glu Ser Gly Ala Gly Lys Thr Glu Asn
223             180              185              190
225 Thr Lys Lys Val Ile Gln Tyr Leu Ala Ser Val Ala Gly Arg Asn Gln
226         195              200              205
228 Ala Asn Gly Ser Gly Val Leu Glu Gln Gln Ile Leu Gln Ala Asn Pro
229     210              215              220
231 Ile Leu Glu Ala Phe Gly Asn Ala Lys Thr Thr Arg Asn Asn Asn Ser
232 225              230              235              240
234 Ser Arg Phe Gly Lys Phe Ile Glu Ile Gln Phe Asn Ser Ala Gly Phe
235             245              250              255
237 Ile Ser Gly Ala Ser Ile Gln Ser Tyr Leu Leu Glu Lys Ser Arg Val
238             260              265              270
240 Val Phe Gln Ser Glu Thr Glu Arg Asn Tyr His Ile Phe Tyr Gln Leu
241     275              280              285
243 Leu Ala Gly Ala Thr Ala Glu Glu Lys Lys Ala Leu His Leu Ala Gly
244     290              295              300
246 Pro Glu Ser Phe Asn Tyr Leu Asn Gln Ser Gly Cys Val Asp Ile Lys
247 305              310              315              320
249 Gly Val Ser Asp Ser Glu Glu Phe Lys Ile Thr Arg Gln Ala Met Asp
250             325              330              335
252 Ile Val Gly Phe Ser Gln Glu Glu Gln Met Ser Ile Phe Lys Ile Ile
253         340              345              350
255 Ala Gly Ile Leu His Leu Gly Asn Ile Lys Phe Glu Lys Gly Ala Gly
256     355              360              365
258 Glu Gly Ala Val Leu Lys Asp Lys Thr Ala Leu Asn Ala Ala Ser Thr
259     370              375              380
261 Val Phe Gly Val Asn Pro Ser Val Leu Glu Lys Ala Leu Met Glu Pro
262 385              390              395              400
264 Arg Ile Leu Ala Gly Arg Asp Leu Val Ala Gln His Leu Asn Val Glu
265             405              410              415
267 Lys Ser Ser Ser Ser Arg Asp Ala Leu Val Lys Ala Leu Tyr Gly Arg
268         420              425              430
270 Leu Phe Leu Trp Leu Val Lys Lys Ile Asn Asn Val Leu Cys Gln Glu
271     435              440              445
273 Arg Lys Ala Tyr Phe Ile Gly Val Leu Asp Ile Ser Gly Phe Glu Ile
274 450              455              460
276 Phe Lys Val Asn Ser Phe Glu Gln Leu Cys Ile Asn Tyr Thr Asn Glu
277 465              470              475              480
279 Lys Leu Gln Gln Phe Phe Asn His His Met Phe Lys Leu Glu Gln Glu
280         485              490              495
282 Glu Tyr Leu Lys Glu Lys Ile Asn Trp Thr Phe Ile Asp Phe Gly Leu
283         500              505              510
```

RAW SEQUENCE LISTING　　　　　　　　DATE: 04/18/2002
PATENT APPLICATION:　US/10/044,303　　TIME: 11:06:15

Input Set : A:\Vectors1.app
Output Set: N:\CRF3\04182002\J044303.raw

```
285 Asp Ser Gln Ala Thr Ile Asp Leu Ile Asp Gly Arg Gln Pro Pro Gly
286         515                 520             525
288 Ile Leu Ala Leu Leu Asp Glu Gln Ser Val Phe Pro Asn Ala Thr Asp
289         530             535             540
291 Asn Thr Leu Ile Thr Lys Leu His Ser His Phe Ser Lys Lys Asn Ala
292 545                 550             555                 560
294 Lys Tyr Glu Glu Pro Arg Phe Ser Lys Thr Glu Phe Gly Val Thr His
295             565             570                 575
297 Tyr Ala Gly Gln Val Met Tyr Glu Ile Gln Asp Trp Leu Glu Lys Asn
298         580             585             590
300 Lys Asp Pro Leu Gln Gln Asp Leu Glu Leu Cys Phe Lys Asp Ser Ser
301         595             600             605
303 Asp Asn Val Val Thr Lys Leu Phe Asn Asp Pro Asn Ile Ala Ser Arg
304     610             615             620
306 Ala Lys Lys Gly Ala Asn Phe Ile Thr Val Ala Ala Gln Tyr Lys Glu
307 625             630             635                 640
309 Gln Leu Ala Ser Leu Met Ala Thr Leu Glu Thr Thr Asn Pro His Phe
310             645             650             655
312 Val Arg Cys Ile Ile Pro Asn Asn Lys Gln Leu Pro Ala Lys Leu Glu
313         660             665             670
315 Asp Lys Val Val Leu Asp Gln Leu Arg Cys Asn Gly Val Leu Glu Gly
316         675             680             685
318 Ile Arg Ile Thr Arg Lys Gly Phe Pro Asn Arg Ile Ile Tyr Ala Asp
319     690             695             700
321 Phe Val Lys Arg Tyr Tyr Leu Leu Ala Pro Asn Val Pro Arg Asp Ala
322 705             710             715                 720
324 Glu Asp Ser Gln Lys Ala Thr Asp Ala Val Leu Lys His Leu Asn Ile
325             725             730             735
327 Asp Pro Glu Gln Tyr Arg Phe Gly Ile Thr Lys Ile Phe Phe Arg Ala
328         740             745             750
330 Gly Gln Leu Ala Arg Ile Glu Glu Ala Arg Glu Gln Arg Leu Gly Ser
331         755             760             765
333 Glu Gln Thr Lys Ser Asp Tyr Leu Lys Arg Ala Asn Glu Leu Val Gln
334     770             775             780
336 Trp Ile Asn Asp Lys Gln Ala Ser Leu Glu Ser Arg Asp Phe Gly Asp
337 785             790             795                 800
339 Ser Ile Glu Ser Val Gln Ser Phe Met Asn Ala His Lys Glu Tyr Lys
340             805             810             815
342 Lys Thr Glu Lys Pro Pro Lys Gly Gln Glu Val Ser Glu Leu Glu Ala
343         820             825             830
345 Ile Tyr Asn Ser Leu Gln Thr Lys Leu Arg Leu Ile Lys Arg Glu Pro
346         835             840             845
348 Phe Val Ala Pro Ala Gly Leu Thr Pro Asn Glu Ile Asp Ser Thr Trp
349     850             855             860
351 Ser Ala Leu Glu Lys Ala Glu Gln Glu His Ala Glu Ala Leu Arg Ile
352 865             870             875                 880
354 Glu Leu Lys Arg Gln Lys Lys Ile Ala Val Leu Leu Gln Lys Tyr Asn
355             885             890             895
357 Arg Ile Leu Lys Lys Leu Glu Asn Trp Ala Thr Thr Lys Ser Val Tyr
```

VERIFICATION SUMMARY　　　　　　　　DATE: 04/18/2002
PATENT APPLICATION: **US/10/044,303**　　　TIME: 11:06:16

Input Set : **A:\Vectorsl.app**
Output Set: **N:\CRF3\04182002\J044303.raw**

```
L:10 M:270 C: Current Application Number differs, Replaced Current Application Number
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:424 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:3
L:424 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:3
L:424 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
```